Available online at www.sciencedirect.com

SciVerse ScienceDirect

Energy Procedia 37 (2013) 2794 – 2801

Energy Procedia

GHGT-11

Common Themes in Risk Evaluation Among Eight Geosequestration Projects

Ken Hnottavange-Telleen[a] *

[a]*Schlumberger Carbon Services, One Hampshire Street, Cambridge, MA 02139, USA*

Abstract

Risk-evaluation workshops were held during early stages of eight carbon dioxide (CO_2) geosequestration (GS) projects. For all projects, the primary entities evaluated were Features, Events, and Processes (FEPs), which are broad concepts pre-selected for their relevance to GS and to each project. Scenarios were later evaluated in some projects, but this paper addresses only the FEPs data.

Design features for all risk workshops included information sharing, participation of 15 to 30 project experts and stakeholders, defined risk receptors (project values), categorical five-point scales of Severity and Likelihood (whose product is the quantity "risk"), and personal expertise ratings. Project values included health and safety of workers, public, and the environment, and usually include budget and schedule, contribution to research goals, and contribution to creating a successful GS industry. A fundamental result for each project was a FEPs ranking in terms of risk to project values.

Workshops are not laboratories, and their characteristics may contain important variables. Nevertheless, their quantitative data enable comparing projects on a risk basis, comparing evaluations from experts and non-experts, and seeking commonalities among higher-risk areas in multiple projects. Data subsets are compared based on (1) the density of (Severity, Likelihood) value pairs among the 25 grid cells of a risk matrix, and (2) a cumulative plot of risk from the lowest-risk to the highest-risk FEPs.

Data subsets from experts, non-experts, and each separate project show a concentration of risk values at low to moderate levels. Relative to experts, non-experts' evaluations are slightly concentrated *toward medium likelihood* and *away from medium severity*. Among the eight projects, some show a greater incidence of low- to moderate-risk FEPs among their total FEPs lists. While it is these apparently lower-risk projects that have been more successful so far, this paper does not attempt to account for the potential effect of risk-evaluation thoroughness.

* Corresponding author. Tel.: +1-508-395-2730.
*E-mail address:*hnottavange-telleen@slb.com.

Selection and/or peer-review under responsibility of GHGT
doi:10.1016/j.egypro.2013.06.164

In each project, some of the highest risks relate to managing a complex of activities and organizations, and to obtaining legal permissions and social license. This suggests that the ability of a GS industry to abate greenhouse gases most depends on addressing *programmatic* risks. Among physical hazards, those typically judged highest are related to driving, construction, and well drilling, rather than to CO_2 itself. Regarding storage characteristics, for the eight selected project sites (evaluated from regional geology and prior to injection-well drilling) the adequacy of injectivity has been judged to pose greater risk than adequacy of capacity.

Selection and/or peer-review under responsibility of GHGT

Keywords: risk; evaluation; comparison; expert judgment

1. Introduction

From 2008 through 2011, risk-evaluation workshops were held during start-up of eight CO_2 geosequestration (GS) projects in North America.

Over this time, in 8 project workshops, 214 scientists and other project professionals …
- Evaluated 741 Features, Events, and Processes (FEPs),
- Completed hundreds of post-workshop spreadsheets,
- Provided 50,427 numerical data points including 10,607 values of "Risk", and
- Wrote 29,734 words of text: 1100 scenarios and 1400 risk treatments.

Workshops are not laboratories and project professionals are not measuring devices that can be precisely calibrated, but conditions, data-collection methods, and evaluation scales were as consistent as practical. The aim of this paper is to investigate whether, given more than 50,000 numerical data points, anything is to be learned from summarizing data across these eight projects. Is it plausible to draw conclusions from comparing risk data across the projects?

Each risk workshop shared many FEPs with other workshops, and some FEPs were identical in wording across all projects and workshops. But some risk areas critical to one project are irrelevant to another, and some FEPs unavoidably take on different meanings in different contexts. Because of this, FEP evaluation is unavoidably entangled with semantics, and FEP risk levels can be compared across projects only with caution. In this paper, higher-risk FEP topics are summarized by group (Table 1), but the main conclusions focus on the distribution of higher and lower risks on an identical evaluation matrix (Figure 1). This spatial distribution is similar for each project, such that comparisons can be made.

The limited nature of inter-project conclusions honors the proper application of project risk evaluation, which is to identify the risk areas that *for a specific project* present the highest risks, and thereby to provide an actionable basis for project risk treatment and risk management. Experience with GS projects may someday provide statistical basis for more "calibrated and quantitative" comparisons of risk levels.

Nomenclature

FEP	Feature, Event, or Process – a conceptual element of a project; compare to *scenario*.
Scenario	A hypothetical but well-defined chain of events. Both FEPs and scenarios can be evaluated in terms of their risks to project values.
Risk	An event, action, or lack of action that has potential negative consequences for a *project value*. Risk is semi-quantified as the product of **Severity** (S) and **Likelihood** (L) of negative impact.
Project Value	An entity or concept identified by the project owner that should be protected from risk. Project values usually include defined achievement objectives, health-safety-environmental standards, and social goals.
Severity	The intensity of negative impact, evaluated according to a categorical scale for each project value.
Likelihood	The estimated, categorically scaled probability that negative impact having a defined severity will occur, given the scope of the risk evaluation.

1.1. Quantifying risk in a new type of project

CO_2 geosequestration (GS) technology is a new feature of the industrial landscape, having almost no project histories that extend back more than a few years. Although GS projects involve mostly familiar industrial activities, they are always undertaken with incomplete knowledge of the earth itself, and they integrate physical components and human organizations that have little history of working together. Because of this, the challenges to project success cannot be comprehensively identified, much less evaluated, by reference to actuarial tables of known event types.

To reduce the uncertainties and to intelligently identify, evaluate, and manage the risks in these projects, Schlumberger Carbon Services in North America has taken an approach to risk evaluation beginning with FEPs [1]. The FEPs approach is conceptual and broad, and it tends toward built-in redundancy that is thought to minimize the chance of overlooking an important risk. FEPs evaluation lays the foundation for developing specific risk-bearing scenarios, whose potential negative impacts can then be countered through project actions and design features. Details of FEPs-based risk evaluation in a GS project can be found in [2], and will be briefly stated here.

Features are static characteristics of a project or its environment, such as stratigraphic layering, faults, existing infrastructure, and installed hardware. *Events* are occurrences such as storms, earthquakes, and startups and shutdowns of CO_2 injection and withdrawal. *Processes* include physical, chemical, and operational aspects like flow of variable-density fluids, well drilling, permitting, and chemical changes.

The “risk” associated with a FEP (or a scenario) is evaluated according to 5-point scales of Severity (S) and Likelihood (L) of potential negative impact to project values. Each FEP’s (S,L) coordinates can be plotted on a risk matrix like Figure 1. The product S×L – the quantity called “Risk” – is therefore scaled from -1 to -25. This value is used for ranking risks, and the risk ranking informs the project’s priorities for developing specific scenarios and undertaking risk treatments.

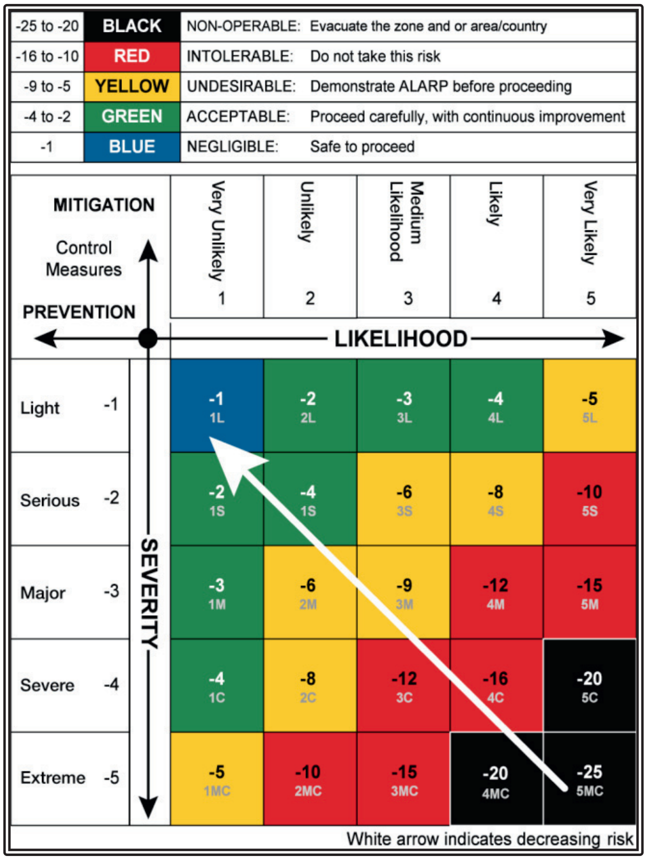

Fig. 1, Risk matrix. The lowest-risk FEPs plot in the blue square at upper left.
Table at top gives project owner's policy guidance (after Schlumberger Standard QHSE-S020).

1.2. Risk-evaluation data from multiple projects

For each of the eight projects, a risk workshop was held early in the organizational project phase. In each workshop, about 90 FEPs (range: 51-137) were evaluated for risk. When the FEPs are plotted on the risk matrix according to their project-average (S,L) coordinates, the FEP population of each matrix cell can be tallied.

The aggregate cell density from the eight projects (Figure 2) can then be used as a standard for comparing risk-evaluation responses between different groups of evaluators (Figure 3) and among different individual projects (Figures 4a, 4b, 4c). The general pattern shown in Figure 2 – many risks evaluated as (-2,2) and progressively fewer risks in cells more distant from (-2,2) – is evident in each large subset of the data that has been examined so far.

For seven of the projects, workshop participants (mainly project professionals, plus a small number of public stakeholders) self-rated their expertise in 8 to 12 technical and project-knowledge areas that were correlated against the risks that they evaluated. In every project, FEP risk rankings computed from values given by the experts are very similar to FEP risk rankings computed from values given by the non-experts who were involved in the same workshop discussions. In detail, the eight-project dataset shows some differences between the (S,L) values given by experts and those given by non-experts (Figure 3).

All-Project Cell Density

Percent of scores per cell

SEVERITY >> \ LIKELIHOOD >>	1	2	3	4	5
-1	10.2%	5.2%	1.7%	0.8%	0.2%
-2	8.7%	**20.9%**	9.9%	2.7%	0.5%
-3	3.2%	9.8%	11.7%	4.3%	0.6%
-4	1.4%	1.8%	2.6%	2.1%	0.4%
-5	0.4%	0.3%	0.2%	0.2%	0.2%

Fig. 3, Risk-matrix cell density. Across all projects and all workshop participants, about one-fifth (20.9%) of FEPs is evaluated as (S=-2, L=2). Each project's individual dataset, and datasets from both "expert" and "non-expert" project professionals show this same density pattern.

Ratio: Non-Expert / Expert Ratings

Project data: 11,627 S,L values

SEVERITY >> \ LIKELIHOOD >>	1	2	3	4	5
-1	1.01	0.95	1.10	0.89	1.08
-2	1.12	1.10	0.99	0.71	0.74
-3	1.06	1.11	0.99	0.79	0.59
-4	0.89	1.09	1.08	0.52	0.68
-5	1.13	0.69	1.76	1.01	0.33

Fig. 2, Differences between project professionals who are expert and those *not* expert in the specific FEP subject matter. Non-experts gave fewer values of L=4 and L=5 (blue cells), and gave 1.76 times the Experts' fraction of (S=-5, L=3) (orange cell).

"Accuracy" of risk evaluation cannot be objectively determined except in hindsight. Greater *certainty* of experts about Likelihood might be inferred, but certainty and accuracy are not necessarily correlated.

Using the same all-project cell density matrix (Fig. 2) for comparison, three individual projects show distinctive patterns of difference (Figs. 4a, b, c). Relative to the eight-project average, Project "A" shows greater cell density (orange) in the cells representing low to moderate risk. Project "B" appears dominated by a small number of very high risks (especially S=-5, L=5). Project "C" has higher density in a large number of the higher-risk cells. Cell density plots do not indicate any specific risk threshold, and they do not inform about which risks can be overcome through suitable treatments. But as such plots accumulate for many projects, they may provide insight into the nature of risk for an individual project, *and/or* the degree of thoroughness of the risk evaluation. An objective measure of one or both of these properties would indeed be useful!

Per-Project Cell Density Relative To All-Project Average

PROJECT "A": Low to moderate risk

SEVERITY >> \ LIKELIHOOD >>	1	2	3	4	5
-1	0.61	0.76	1.06	0.99	0.30
-2	1.19	1.17	1.26	0.91	0.30
-3	0.86	1.26	1.14	0.66	0.13
-4	0.83	0.85	0.81	0.26	0.35
-5	0.53	0.23	0.46	0	0

PROJECT "B": Dominant very high risks

SEVERITY >> \ LIKELIHOOD >>	1	2	3	4	5
-1	0.88	0.79	0.82	0.49	0.27
-2	1.24	1.09	0.89	0.95	1.12
-3	0.93	1.14	0.87	1.13	1.26
-4	0.71	0.53	0.86	1.17	2.09
-5	1.61	0.43	1.25	1.05	5.58

PROJECT "C": Many higher risks

SEVERITY >> \ LIKELIHOOD >>	1	2	3	4	5
-1	0.93	0.92	1.45	1.42	0.58
-2	0.71	0.86	1.07	1.23	1.00
-3	0.71	0.88	1.10	1.30	1.49
-4	0.44	1.22	1.62	1.73	1.16
-5	0.81	1.70	2.41	2.45	0.75

Fig. 4, Relative to all-project cell density (Figure 2), three individual projects show distinctive cell-density patterns. The patterns may signal each project's risk characteristics, but are also probably influenced by the thoroughness of the risk evaluation itself.

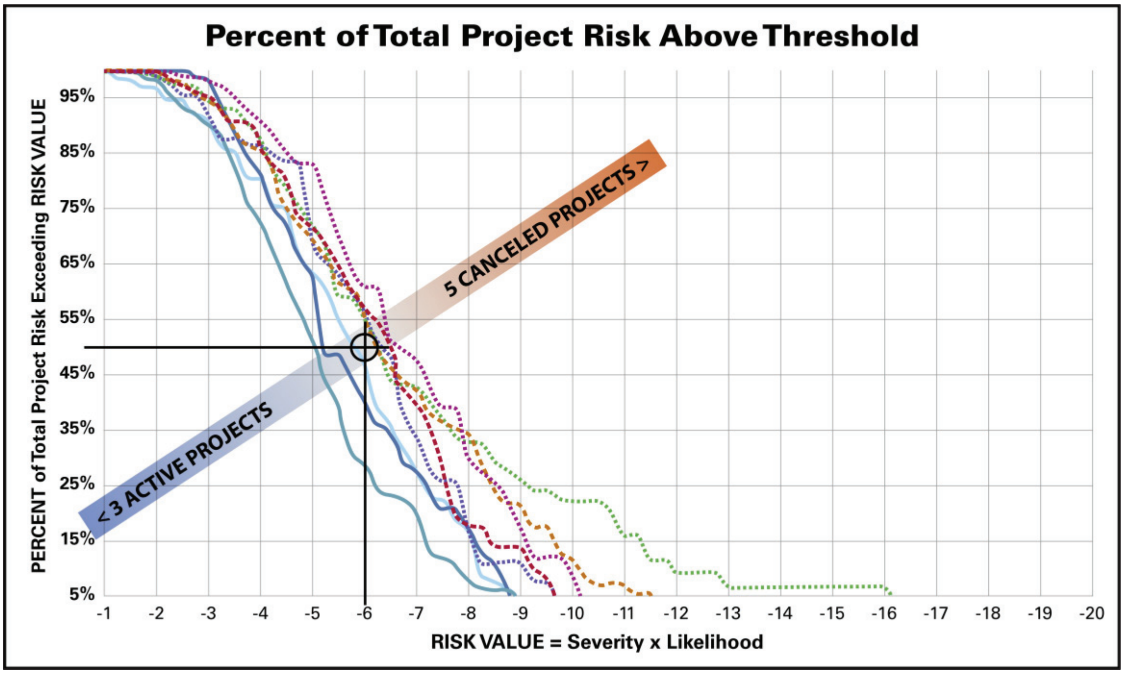

Fig. 5, Risk was evaluated early in each project. In 3 projects (solid lines), more than half of total project risk existed at risk values of **-1 through -6**. In 5 projects, more than half of total project risk existed at values of **-6 and greater.** These higher-risk projects were later canceled.

1.3. Hints of an objective measure of project risk?

After ranking, a project's risk values can be cumulated from lowest to highest. In Figure 5, 100% of the project's total risk exists above the lowest risk level (R=-1), while 0% remains above the highest risk level (R=-25). Among the eight projects, the five that have their greatest fractions of risk remaining above the risk value of -6 have been canceled, while the three with their greatest risk fractions at lower risk levels remain active. That is, the cumulative risk line appears to have separated the "active" from the "canceled" projects. Does this observation have predictive value for a project, or design value for a risk workshop? Again, more study is needed. For the inactive projects, the causes of cancellation have usually been identified during the initial risk workshop, and the causes have typically been organizational or economic rather than geotechnical, engineering, or hazard-related.

1.4. Common themes in GS project risk

Risk is project-specific and site-specific, but there may be recurrent themes. Table 1 cites common threads among the FEPs that have frequently been evaluated as presenting higher risk to project values. The source material for the table includes all FEPs ranked #1-#50 overall (out of 741 FEPs total), plus additional FEPs ranked #1-#10 for individual projects only. Among FEPs ranked as low risk there is less commonality.

Table 1, Higher-risk themes in CCS projects. Within the 8 projects, FEPs evaluated as higher-risk have been grouped according to topic. The 12 listed FEP topic groups indicate themes commonly found to present higher risks to project success.

FEP TOPIC GROUP	How many FEPs in topic group were called "high risk"?
Permissions: Regulatory	8
Project organization and management	6
Economics: Front-end funding and costs	5
Operator Liability	5
Geology of reservoir	5
Project organization and management: Procurement	5
Economics: Operating cash flow	4
Community	4
CO_2 capture & delivery design	3
Permissions: Property rights	3
CO_2 delivery system operation	2
Project organization and management: Contracting	2

1.5. Severity vs. Likelihood as a key to risk treatments

Severity and Likelihood are the components used to quantify risk as S×L. Their relative contribution to risk can also point toward the more efficient risk-reduction strategy. Among the 10,607 (S,L) pairs there are nearly 7000 pairs where $S \neq L$ (that is, the absolute value of "S", ignoring the "-" sign). At low risk,

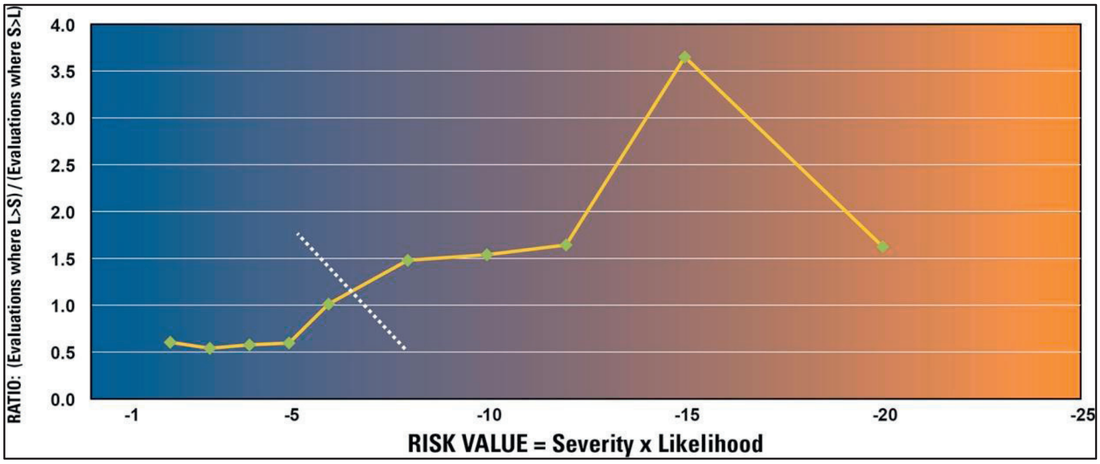

Fig. 6, Prevention vs. Mitigation. For a given risk level (x-axis), points show the ratio of the number of (S,L) pairs in which L>S to the number of pairs in which S>L. Pairs in which S=L are excluded. At higher risk, most pairs have L>S.

most pairs have S>L; Severity dominates (Figure 6). At higher risk, most pairs have L>S; Likelihood dominates. Of course, specific risks demand specific responses and some risks are best addressed through severity reduction (mitigation) … but overall, the first treatment strategy in risk management is PREVENTION.

1.6. Conclusions

We cannot measure risk with calibrated laboratory instruments. And those of us who evaluate risk – even the subject-matter experts! – bring many influences besides objective datasets to the workshop [3]. But we can design risk-evaluation workshops to *benefit* from the different points of view and sources of information available to participants who have different experience [4], [5]. If we can collect risk-evaluation data in ways that are consistent, yet that do not over-determine outcomes and simply reinforce bias, maybe we can objectively extract important risk themes within single projects; objectively compare risks across projects; and – most important – protect projects from the sharpest bites of Taleb's black swan [6].

Acknowledgements

The author gratefully acknowledges the efforts of the more than 150 scientists, managers, scientist-managers, engineers, attorneys, project neighbors, and others, known only unto themselves and a large database, who have taken part in the grueling risk workshops whose data is reported in this paper.

References

[1] Quintessa Ltd. Generic CO_2 FEP Database, Version 1.1.0. Open access on-line database. http://www.quintessa.org/co2fepdb/

[2] Hnottavange-Telleen K et al, 2011, Risk management in a large-scale CO2 geosequestration pilot project, Illinois, USA; Energy Procedia Vol. 4 (Elsevier), 10th Intl Conference on Greenhouse Gas Control Technologies, Amsterdam NL

[3] Kahneman D, Slovic P, Tversky A (editors). Judgment under uncertainty: heuristics and biases. Cambridge, UK: Cambridge University Press; 1982.

[4] Surowiecki J. The wisdom of crowds: Why the many are smarter than the few and how collective wisdom shapes business, economies, societies and nations. New York: Little, Brown; 2004.

[5] Yousuf MI (2007). Using experts' opinions through Delphi technique. Practical Assessment Research & Evaluation, 12(4).

[6] Taleb NN. The black swan: The impact of the highly improbable. New York: Random House; 2007.